A MODIFIED COMPUTATIONAL MODEL OF ANT COLONY SYSTEM IN DNA SEQUENCE DESIGN

SERI MASTURA MUSTAZA

UNIVERSITI TEKNOLOGI MALAYSIA

A MODIFIED COMPUTATIONAL MODEL OF ANT COLONY SYSTEM IN DNA SEQUENCE DESIGN

SERI MASTURA MUSTAZA

A project report submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Engineering (Mechatronics & Automatic Control)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JANUARY 2012

ACKNOWLEDGEMENT

First and foremost, the deepest gratitude of all shall be bestowed to Allah the Almighty and The Merciful for all the insight which He gave to us that lead to the completion of this project.

I would like to express my sincere gratitude to my supervisor Dr. Zuwairie Ibrahim for his supervision, patience, motivation, and immense knowledge which have guided me throughout the process of completing this research. His enthusiasm in the research area is very inspiring and has definitely enriched my growth as a student and researcher. Without his knowledge and encouragement, this study would not have been successful.

I convey a special thank you to my fellow friends especially Amar Faiz Zainal Abidin, Mohd Hafiz Othman and Lai YeeYang for the knowledge that was shared, the time the was spent and support that was provided to complete all the work and meet all the deadline throughout my years in UTM. It would not have been the same without you guys.

I would also like to thank my mum for always stood by me, giving me strength and praying for my success, as well as my dad, my brother and my sisters who always encourage me with their best wishes and prayers.

Last but not least, I would like to thank my husband whose love, support and persistent confidence in me, has tremendously help me to finish my study.

ABSTRACT

Major principle behind the development of computational intelligence is to address complex problem of real world application. Over the years, numerous computational intelligence algorithms have been developed in finding a solution to combinatorial optimization problem. Ant colony system (ACS) algorithm is one of the biologically inspired algorithms that have been applied to effectively solve various combinatorial optimization problems. In this study, ACS is going to be employed in solving DNA sequence design which is a study under the topics of DNA computing. The dependability of DNA computation is highly influenced by the information represents on the DNA strand and the strand reaction. We desire a set of stable double stranded DNA to retrieve the information encoded on the DNA sequence and to operate the computation without output error. To accomplish this, the DNA sequence design problem requires a set of objectives to be optimized and some constraints to be fulfilled. Therefore, DNA sequence design can be regarded as a constrained multi-objectives design problem. The multi-objective design problem is simplified into single-objective using the weighted sum method and objective functions used to obtain a good DNA sequence are $H_{measure}$, *similarity*, *hairpin*, and *continuity*. The sequence is subjected to two constraints which are T_m and $GC_{content.}$ The problem is modeled using finite state machine where each node represents the DNA bases {*A, C, T, G*}. In this study, 9 sets of studies have been conducted using 5, 7, 10, 15, 20, 25, 30, 35 and 40 agents/ants each with 100 independent runs. The number of iterations is set to be 300 for each set. Observation and analysis of the model with increasing number of ants was made and the performance of the model is measured by comparing the result with existing algorithm such as Genetic Algorithm (GA), Multi-Objective Evolutionary Algorithm (MOEA), Particle Swarm Optimization (PSO) etc. Based on the result, the suitable number of ants used for DNA sequence design was also proposed.

ABSTRAK

Prinsip utama di sebalik pembangunan pengkomputeran pintar adalah untuk menangani masalah kompleks yang melibatkan aplikasi dunia sebenar. Sejak kebelakangan ini, pelbagai algoritma serta perisian penkomputeran pintar telah dibangunkan dalam mencari penyelesaian kepada masalah pengoptimuman kombinatorik. Algoritma *Ant Colony System* (ACS) adalah salah satu algoritma yang telah digunakan dengan berkesan dalam menyelesaikan pelbagai masalah pengoptimuman kombinatorik. Dalam kajian ini, algoritma ACS telah digunakan dalam menyelesaikan masalah rekabentuk turutan DNA. Kebolehpercayaan pengkomputeran DNA sangat dipengaruhi oleh maklumat yang terdapat pada lembar DNA serta tindak balas antara DNA. Set DNA yang stabil adalah sangat diperlukan bagi mendapatkan maklumat yang tepat dan memastikan pengendalian pengiraan tanpa ralat. Untuk mencapai tujuan ini, masalah reka bentuk jujukan DNA memerlukan satu set objektif yang perlu dioptimumkan dan beberapa kekangan yang perlu dipenuhi. Oleh itu, masalah turutan DNA boleh dianggap sebagai masalah rekabentuk multi-objektif dan telah dipermudahkan menjadi masalah satu-objektif menggunakan kaedah jumlah wajaran. Fungsi objektif yang digunakan bagi mendapatkan turutan DNA yang baik adalah $H_{measure}$, *similarity*, *hairpin*, and *continuity* dan tertakluk kepada dua kekangan iaitu T_m and $GC_{content}$. Masalah ini dimodel menggunakan mesin keaadan terhingga dimana setiap nodus mewakili asas DNA {*A, C, T, G*}. Dalam kajian ini, 9 set kajian telah dijalankan menggunakan 5, 7, 10, 15, 20, 25, 30, 35 dan 40 bilangan agen/semut. Pemerhatian dan analisis dengan peningkatan bilangan agen telah dibuat serta prestasi model diukur melalui perbandingan dengan algoritma yang sedia ada seperti as *Genetic Algorithm* (GA), *Multi-Objective Evolutionary Algorithm* (MOEA), *Particle Swarm Optimization* (PSO) dan lain-lain. Hasil kajian ini juga digunakan bagi mencadangkan bilangan agen/semut yang sesuai bagi aplikasi masalah rekabentuk turutan DNA.